

# 深度学习在湿度环境量预测中的研究

张丽梅<sup>1</sup>, 蔡定霄<sup>1</sup>, 郑爱军<sup>2</sup>, 肖贵友<sup>2</sup>, 夏泽勇<sup>2</sup>

(1. 中国电建集团昆明勘测设计研究院有限公司, 云南 昆明 650051;

2. 西藏开发投资集团有限公司, 西藏 拉萨 850000)

**摘要:** 为了构建适用于金沙江中游干热河谷地区的湿度高精度预测方法, 以该地区近7年的逐日湿度观测数据为研究对象, 采用深度学习技术构建预测模型, 研究将80%的时序数据作为训练集, 20%作为测试集, 通过决定系数( $R^2$ )、平均绝对误差(MAE)和均方根误差(RMSE)评估模型性能, 重点对比分析了RNN、GRU、BiLSTM和BiLSTM-Attention四种模型的预测效果。结果表明: 基础RNN和GRU模型的训练集 $R^2$ 分别为0.9227和0.8414, 预测精度有限; 而BiLSTM模型展现出显著优势, 训练集和测试集的 $R^2$ 分别达到0.9796和0.9770, MAE降至1.7257% RH和2.1389% RH。引入注意力机制后, BiLSTM-Attention模型性能进一步提升, 训练集和测试集 $R^2$ 分别提高至0.9875和0.9821, 预测误差进一步降低。研究证实BiLSTM及其注意力机制改进模型不仅具有优异的训练数据拟合能力, 更能有效预测未知数据, 展现出强大的泛化性能。研究还基于Scott核密度估计构建了概率预测框架, 为湿度变化提供了可靠的置信区间估计。

**关键词:** 湿度; 数据驱动; 深度学习; 预测模型; 泛化性能

中图分类号: TV641

文献标识码: A

文章编号: 1672-1144(2025)06-0192-08

## Deep Learning in Humidity Environmental Prediction

ZHANG Limei<sup>1</sup>, CAI Dingxiao<sup>1</sup>, ZHENG Aijun<sup>2</sup>, XIAO Guiyou<sup>2</sup>, XIA Zeyong<sup>2</sup>

(1. Hydrochina Kunming Engineering Corporation Limited, Kunming, Yunnan 650051, China;

2. Xizang Development and Investment Group Co., Ltd., Lhasa, Tibet 850000, China)

**Abstract:** In order to construct a high-precision prediction method for humidity applicable to the middle reaches of the Jinsha River dry and hot valley, the daily humidity observation data the last 7 years in this area were taken as the research object, and the prediction model was constructed by using deep learning technology. The research allocated 80% of the time-series data as the training set and 20% as the test set, evaluating model performance using the coefficient of determination ( $R^2$ ), mean absolute error (MAE), and root mean square error (RMSE). A comparative analysis was conducted on the predictive efficacy of four models: RNN, GRU, BiLSTM, and BiLSTM-Attention. The results indicate that the basic RNN and GRU models achieved training set  $R^2$  values of 0.9227 and 0.8414, respectively, demonstrating limited predictive accuracy. In contrast, the BiLSTM model exhibited significant advantages, with training and test set  $R^2$  values reaching 0.9796 and 0.9770, respectively, while MAE decreased to 1.7257% RH and 2.1389% RH. The introduction of an attention mechanism further enhanced the performance of the BiLSTM-Attention model, increasing training and test set  $R^2$  values to 0.9875 and 0.9821, respectively, and further reducing prediction errors. The study confirms that the BiLSTM model and its attention-enhanced variant not only demonstrate excellent fitting capability on training data but also effectively predict unknown data, showcasing strong generalization performance. Additionally, a probabilistic prediction framework based on Scott's kernel density estimation was developed, providing reliable confidence interval estimates for humidity variations.

**Keywords:** humidity; data-driven; deep learning; predictive models; generalization performance

机器学习在诸多领域取得较大成就,在环境量领域成果也很显著。深度学习是机器学习的一个新领域,由 Hinton 等于 2006 年提出,它被引入机器学习使其更接近于最初的目标——人工智能。深度学习是一种基于人工神经网络的机器学习方法,通过学习样本数据的内在规律和表示层次,让机器能够像人一样具有分析学习能力,识别文字、图像和声音等数据。

已有文献表明,机器学习已取得较为丰硕的成果:薛义宽等<sup>[1]</sup>提出了基于循环神经网络(recurrent neural network, RNN)的 STAT3 抑制剂的设计方法,吴昊等<sup>[2]</sup>提出了基于自适应循环神经网络的异常用电行为检测方法,游弈擎等<sup>[3]</sup>提出了基于循环神经网络的快速预测标准单元延时方法,王亚茹等<sup>[4]</sup>提出了基于 GRU 的质量聚合模块实现整张图像的质量分数,宁悦等<sup>[5]</sup>通过融合门控循环单元模型和注意力机制,提高了 U 型槽微带贴片天线设计精准度和速度,柳天虹等<sup>[6]</sup>提出了一种基于分位数回归的改进权重 GRU 风电功率区间预测模型(QR-EGRU),刘翌晗等<sup>[7]</sup>提出了一种分布式非线性映射和并行输入的 BiLSTM 高预测精度的软测量模型,李鑫帅等<sup>[8]</sup>提出了基于时序模型用于猕猴桃果园果实膨大期土壤水分精准监测的 BiLSTM 模型,畅晗等<sup>[9]</sup>提出了一种基于浣熊优化算法(COA)优化变分模态分解(VMD)算法与融合卷积神经网络(CNN),双向长短期记忆网络(BiLSTM)和注意力机制的出口 SO<sub>2</sub> 浓度预测模型,赵法森等<sup>[10]</sup>提出了可实现煤矿事故隐患排查相关信息系统提供快速录入的 BiLSTM + Attention 模型。

机器学习在环境量研究中也取得了一定研究成果,秦一菲等<sup>[11]</sup>基于 XGBoost 算法构建了储粮湿度预测模型;罗凤春等<sup>[12]</sup>利用支持向量机(SVM)建立了多层土壤湿度模拟模型;张疏桐等<sup>[13]</sup>引入 Kolmogorov-Arnold 网络(KAN),与随机森林(RF)、循环神经网络(RNN)和长短期记忆网络(LSTM)3 种机器学习算法比较,确定了不同时间步长下温湿度预测性能最优模型;毛晓娟等<sup>[14]</sup>提出了准确地预测未来温室内温度变化的 GWO-LSTM 模型;智协飞等<sup>[15]</sup>利用 LSTM、NN、BREM 和 SUP 等深度学习方法实现了中国地面气温的多模式集成预报;张观山等<sup>[16]</sup>提出了基于 LSTM-AT 模型的温室空气温度预测。

综上,在已有的研究成果中机器学习尤其是深度学习在医学、图像识别、电工学、农业等领域取得

了一定成果。在环境量预测方面,虽然对温度、湿度等单一气象要素已有部分研究,但针对长时间序列空气湿度这一关键环境参数的预测研究仍属空白。湿度作为影响生态环境、农业生产、工业制造和人体健康的核心环境因子,其精准预测具有重要的实践价值:在农业领域可为干热河谷地区作物灌溉和病虫害防治提供决策依据;在电力行业能有效预防设备结露故障;在健康医疗领域有助于呼吸道疾病预防;在建筑领域可优化室内环境控制系统。

为此,基于金沙江中游干热河谷地区接近 7 年湿度逐日实测值,构建 RNN、GRU、BiLSTM、BiLSTM-Attention 等深度学习模型进行训练,采用湿度逐日实测值的前 80% 时间序列数据作为训练集,剩余 20% 作为测试集,通过决定系数( $R^2$ )、平均绝对误差(MAE)、均方根误差(RMSE)评判模型对训练数据的拟合程度和在对测试集中未知数据的预测能力,验证模型的泛化性,研究适用于本项目的湿度预测的深度学习方法。考虑到湿度受复杂地形和强烈蒸发作用,变化具有高度非线性与不确定性,进一步基于 Scott 核密度估计的概率预测框架进行了区间预测成果分析。本研究建立的预测框架,为复杂气候区环境量预测提供了可迁移的技术方案,对实现生态环境智能监测具有重要意义。

## 1 硬件设备及深度学习方法

### 1.1 硬件设备

为对不同深度学习方法计算效率进行分析,四种不同深度学习方法采用同一台计算机 T 490 笔记本进行计算。

### 1.2 采用的深度学习方法

#### 1.2.1 深度学习模型特性介绍

循环神经网络(RNN)作为一种反馈神经网络,在处理时间序列数据上比传统的前馈网络更能体现数据在时间上的依赖关系<sup>[17]</sup>,其核心思想是通过内部的记忆机制捕捉时间序列中的时序信息。循环神经网络具有循环结构,能够有效建模输入数据之间的动态时间关系。该模型在每个时间步共享相同的参数,这种权重共享机制显著降低了模型的复杂度和训练参数数量。此外,循环神经网络能够记忆历史信息,并在后续时间步中利用这些信息进行预测或决策。由于参数在时间轴上共享,循环神经网络可以灵活处理不同长度的输入序列,实现信息在序列内部的流动。然而,循环神经网络也存在一些固有缺陷,最突出的问题是梯度消失或爆炸现象,这导

致模型难以捕捉长期依赖关系,只能有效利用较短上下文信息。随着序列长度增加,梯度问题会严重影响模型性能,使训练过程不稳定且难以收敛。在计算效率方面,循环神经网络需要按时间步依次展开计算,这种串行特性使得处理长序列时效率较低。同时,计算过程的时序依赖性限制了模型的并行化能力,导致训练速度较慢。这些因素共同导致循环神经网络在实际应用中面临较大的训练难度。

为解决传统循环神经网络的梯度消失和梯度爆炸问题,研究者们提出了多种改进架构。1997年, Hochreiter 和 Schmidhuber<sup>[18]</sup> 提出了长短时记忆网络(LSTM),通过引入精密的门控机制(包括输入门、遗忘门和输出门)来调控信息流动,显著提升了模型捕捉长期依赖关系的能力。门控循环单元是循环神经网络的一种变体<sup>[19]</sup>,该模型通过合并部分门控结构(如将输入门和遗忘门整合为更新门),在保持与长短时记忆网络相近性能的同时降低了计算复杂度。

在架构创新方面,双向循环神经网络(Bi-RNN)通过组合前向和后向循环神经网络层,使模型能够同时利用历史和未来的上下文信息。这一思想很快被拓展到长短时记忆网络架构,形成了双向长短时记忆网络(BiLSTM)。

Attention 机制最早在 2014 年由 Bahdanau 等<sup>[20]</sup>提出的,用于改进神经机器翻译(NMT)任务。Attention 机制通过动态计算输入序列各部分的权重,让模型在解码时能够聚焦于相关的输入部分,从而提升翻译质量,核心是计算输入序列(如源语言句子)和输出序列(如目标语言句子)之间的相关性权重,使得模型在生成每个输出词时,能够自适应地关注输入序列中最相关的部分。可以缓解长序列信息丢失问题:相比循环神经网络、长短时记忆网络的固定长度编码,Attention 能动态关注关键信息,有效捕捉长期依赖。

### 1.2.2 循环神经网络(RNN)模型

循环神经网络是一种具有循环结构的神经网络,其核心特点是能够处理序列数据,并捕捉时间序列数据中的动态特征。通过在网络中引入循环结构,使得网络能够保存之前的信息,并在当前时刻的计算中加以利用。循环神经网络的基本结构由循环层构成,每个循环层都有一个权重矩阵用于进行时间序列数据的处理。每个循环层保存着一个隐藏状态(Hidden State),该状态会在下一个循环层中作为输入,并随着时间序列的推进进行更新。这种结构

让循环神经网络能够处理时间序列数据中的依赖关系,并理解序列数据中的时序信息。

假设输入数据  $X$  为  $[n, m]$ , 其中  $n$  表示时间序列数据的长度,  $m$  表示特征维度; 隐藏状态  $H$  为  $[h, m]$ , 其中  $h$  表示隐藏层的大小; 输出  $Y$  为  $[n, 1]$ 。在  $t$  时刻, 输入为  $X_t$ , 上一时刻的隐藏状态为  $H_{t-1}$ , 当前时刻的隐藏状态  $H_t$  和输出  $Y_t$  的计算如下:

隐藏状态更新公式:

$$H_t = f_1(W_{xh}X_t + W_{hh}H_{t-1} + b_h) \quad (1)$$

式中:  $W_{xh}$  为输入层到隐藏层的权重矩阵;  $W_{hh}$  为隐藏层到隐藏层的权重矩阵;  $b_h$  为隐藏层的偏置项;  $f_1$  为隐藏层的激活函数, 常用的激活函数有  $\tanh$  或  $\text{ReLU}$  等。

输出计算公式:

$$Y_t = f_2(W_{hy}H_t + b_y) \quad (2)$$

式中:  $W_{hy}$  为隐藏层到输出层的权重矩阵;  $b_y$  为输出层的偏置项;  $f_2$  为输出层的激活函数, 根据具体任务选择合适的激活函数, 如在分类任务中可能使用  $\text{softmax}$  函数。

其传播分为向前传播和向后传播两种:

#### (1) 前向传播过程

在神经网络的前向传播过程中, 循环神经网络每次处理一个时间步的数据, 并更新隐藏状态。初始时, 隐藏状态  $H_0$  通常初始化为零向量。然后按照时间顺序依次计算每个时间步的隐藏状态和输出:

$$t = 0 \text{ 时}, H_0 = 0$$

$t > 0$  时, 按照上述隐藏状态更新公式和输出计算公式依次计算  $H_t$  和  $Y_t$ 。

#### (2) 反向传播过程

在神经网络的反向传播过程中, 对于每个时间步, 需要计算时刻  $t$  的输出误差  $\delta_t$  和隐藏状态误差  $e_t$ 。然后, 通过反向传播算法计算权重的梯度, 并进行权重更新。

输出误差计算:

$$\delta_t = (Y_t - T_t)f_2'(Y_t) \quad (3)$$

式中:  $T_t$  表示时刻  $t$  的实际输出值;  $f_2'$  是输出层激活函数的导数。

隐藏状态误差计算:

$$e_t = (\delta_t W_{hy})f_1'(H_t) + e_{t+1}W_{hh}f_1'(H_t) \quad (4)$$

式中:  $f_1'$  为隐藏层激活函数的导数。

权重更新公式:

$$W = W - \alpha \Delta W \quad (5)$$

式中:  $\alpha$  为学习率;  $\Delta W$  为权重的梯度。

### 1.2.3 门控循环单元(GRU)模型

门控循环单元是循环神经网络的一种变体,它旨在解决传统循环神经网络在处理长序列时存在的梯度消失和梯度爆炸问题,通过引入门控机制来控制信息的流动,从而有效地捕捉序列数据中的长期依赖关系。

门控循环单元模型包含两个主要的门控机制:更新门(Update Gate)和重置门(Reset Gate),具体原理如下:

更新门:决定当前隐藏状态与之前隐藏状态的相关程度,控制前一时刻的状态信息被代入到当前状态中的程度。更新门 $z_t$ 通过一个Sigmoid函数输出一个范围在 $[0,1]$ 之间的值,计算公式如下:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (6)$$

式中: $W_z$ 为输入 $x_t$ 到更新门的权重矩阵; $U_z$ 为上一时刻隐藏状态 $h_{t-1}$ 到更新门的权重矩阵; $b_z$ 为更新门的偏置项; $\sigma$ 为Sigmoid激活函数,其作用是将输出值压缩到 $[0,1]$ 区间。

重置门:确定了前一时刻的隐藏状态如何影响当前时刻的候选隐藏状态,控制要遗忘多少过去的信息。重置门 $r_t$ 同样通过Sigmoid函数输出一个范围在 $[0,1]$ 之间的值,计算公式如下:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (7)$$

式中: $W_r$ 为输入 $x_t$ 到重置门的权重矩阵; $U_r$ 为上一时刻隐藏状态 $h_{t-1}$ 到重置门的权重矩阵; $b_r$ 为重置门的偏置项。

候选隐藏状态 $\tilde{h}_t$ 结合了当前输入 $x_t$ 和重置门 $r_t$ 的影响,计算公式如下:

$$\tilde{h}_t = \tanh(W x_t + r_t \odot U h_{t-1} + b) \quad (8)$$

式中: $W$ 为输入 $x_t$ 到候选隐藏状态的权重矩阵; $U$ 为上一时刻隐藏状态 $h_{t-1}$ 到候选隐藏状态的权重矩阵; $b$ 为候选隐藏状态的偏置项; $\odot$ 表示逐元素相乘; $\tanh$ 是双曲正切激活函数,其作用是将输出值压缩到 $[-1,1]$ 区间。

根据更新门 $z_t$ 的值和上一时刻的隐藏状态 $h_{t-1}$ 与当前时刻的候选隐藏状态 $\tilde{h}_t$ ,最终计算出当前时刻的隐藏状态 $h_t$ ,计算公式如下:

$$h_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{t-1} \quad (9)$$

当 $z_t$ 接近1时, $h_t$ 主要由 $h_{t-1}$ 决定,即保留了更多的历史信息;当 $z_t$ 接近0时, $h_t$ 主要由 $\tilde{h}_t$ 决定,即更多地考虑了当前输入的信息。

### 1.2.4 双向长短时记忆网络(BiLSTM)模型

双向长短时记忆网络由前向长短时记忆网络和

后向长短时记忆网络组成,通过结合两个方向的隐藏状态来捕捉序列的全局上下文信息。以下是双向长短时记忆网络核心计算步骤和公式:

#### (1) 前向长短时记忆网络(Forward LSTM)

对输入序列 $X = (x_1, x_2, \dots, x_T)$ 按时间步 $t = 1 \rightarrow T$ 计算:

##### ① 输入门(Input Gate)

控制新信息的保留程度:

$$i_t = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + b_i) \quad (10)$$

式中: $W_{xi}$ 为输入到输入门的权重矩阵; $W_{hi}$ 为隐藏状态到输入门的权重矩阵; $b_i$ 为输入门偏置项; $\sigma$ 为sigmoid激活函数。

##### ② 遗忘门(Forget Gate)

控制旧信息的遗忘程度:

$$f_t = \sigma(W_{xf} x_t + W_{hf} h_{t-1} + b_f) \quad (11)$$

式中: $W_{xf}$ 为输入到遗忘门的权重矩阵; $W_{hf}$ 为隐藏状态到遗忘门的权重矩阵; $b_f$ 为遗忘门偏置项。

##### ③ 候选记忆

生成新候选记忆:

$$\tilde{C}_t = \tanh(W_{xc} x_t + W_{hc} h_{t-1} + b_c) \quad (12)$$

式中: $W_{xc}$ 为输入到候选状态的权重矩阵; $W_{hc}$ 为隐藏状态到候选状态的权重矩阵; $b_c$ 为候选状态偏置项。

##### ④ 记忆更新

结合遗忘门和输入门更新记忆:

$$C_t^- = f_t \odot C_{t-1}^- + i_t \odot \tilde{C}_t \quad (13)$$

式中: $\odot$ 为逐原始相乘(Hadamard积)。

##### ⑤ 输出门

控制隐藏状态的输出:

$$o_t = \sigma(W_{xo} x_t + W_{ho} h_{t-1} + b_o) \quad (14)$$

$$h_t^- = o_t \odot \tanh(C_t^-) \quad (15)$$

#### (2) 后向长短时记忆网络计算公式

后向长短时记忆网络的计算与前向对称,只是时间方向相反( $t = T \rightarrow 1$ ):

$$h_t^+ = \text{LSTM}(x_t, h_{t+1}^+) \quad (16)$$

所有参数符号与前向长短时记忆网络相同,但实际训练中是独立的两组参数。

#### (3) 双向状态合并

$$h_t = [h_t^-; h_t^+] \quad (17)$$

### 1.2.5 BiLSTM-Attention模型

Attention属于一种模拟人脑注意力特征的模型<sup>[21]</sup>。单独的双向长短时记忆网络模型将所有信息

输入后直接输出最后一个时刻的隐层状态,而 Attention 将为每个信息分配一个权重,通过多次训练后的评估进行权重的调整。给影响程度大的信息分配高的权重,使模型对权重高的信息更关注,从而减少计算负担,得到较理想的预测精度,其状态转换过程为:

$$e_i = z \tanh(W h_i + b) \quad (18)$$

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^N e_j} \quad (19)$$

$$C_i = \sum_{i=1}^N \alpha_i h_i \quad (20)$$

式中: $e_i$  为时刻注意力概率分布的值; $z$ 、 $W$ 、 $b$  为模型学习参数; $\alpha_i$  为注意力权重; $N$  为模型输入向量的维度; $C_i$  为  $i$  时刻的输出。

表 2 不同深度学习模型评价指标成果

计算模型	训练集			测试集			计算耗时 /s
	$R^2$	MAE/% RH	RMSE/% RH	$R^2$	MAE/% RH	RMSE/% RH	
RNN	0.9227	3.3314	5.7977	0.7130	9.2917	12.3272	162.16
GRU	0.8414	5.9380	8.3051	0.7787	7.8976	10.8246	189.82
BiLSTM	0.9796	1.7257	2.3724	0.9770	2.1389	3.0405	50.00
BiLSTM-Attention	0.9875	1.5140	2.0596	0.9821	1.8295	2.5649	56.20

图 1—图 2 分别给出了循环神经网络深度学习模型在训练集和测试集中实测值和预测值的对比成果,从图 1 中可看出,循环神经网络模型在训练集中预测值反映了实测值的整体趋势,但是在极大值和极小值预测中精度较差,无法真实反应实测值的极值情况;图 2 给出了循环神经网络深度学习模型在未知数据测试集中的拟合情况,显然测试集也无法完全捕捉实测值的极值情况,计算值和实测值整体偏差较大。

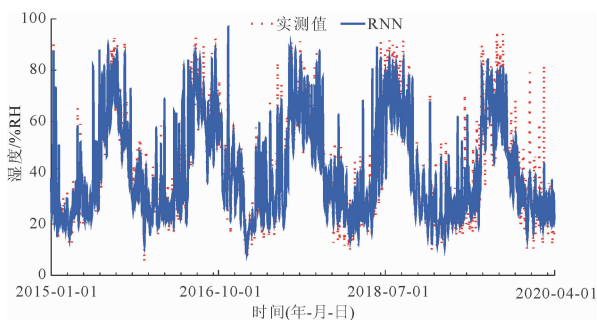


图 1 循环神经网络 (RNN) 深度学习模型训练集计算成果

图 3—图 4 分别给出了门控循环单元深度学习模型训练集和测试集计算成果,从图 3 中可看出,门

## 2 计算结果

### 2.1 四种深度学习模型预测结果分析

为研究 RNN、GRU、BiLSTM、BiLSTM-Attention 四种深度学习模型的拟合程度、预测能力和计算效率,表 2 给出了四种不同深度学习模型的决定系数 ( $R^2$ )、平均绝对误差 (MAE)、均方根误差 (RMSE) 评价指标和计算耗时成果。表 2 可明显看出四种深度学习方法在训练集中决定系数  $R^2$  均大于 0.8,测试集中决定系数  $R^2$  均大于 0.7; MAE、RMSE 评价指标在训练集中分别小于 6% RH、9% RH,在测试集中分别小于 10% RH、13% RH,四种深度模型均可较好地实现了湿度的拟合和预测;计算耗时均在 200 s 之内,说明四种深度学习模型可在短时间内实现深度学习模型的训练和预测,模型计算效率较高。

控循环单元模型在训练集中预测值不能有效捕捉极大值和极小值,无法真实反应实测值的极值情况,与循环神经网络深度学习模型规律相同;图 4 给出了门控循环单元深度学习模型在未知数据测试集中的拟合情况,显然测试集中门控循环单元深度学习模型计算值与实测值偏差仍有一定偏差,但是整体预测成果相对循环神经网络深度学习模型更加接近实测值。

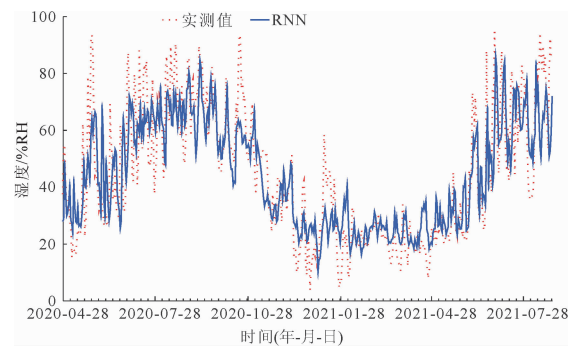


图 2 循环神经网络 (RNN) 深度学习模型测试集计算成果

图 5—图 6 分别给出了双向长短时记忆网络深度学习模型训练集和测试集计算成果,从图 5 中可看出,双向长短时记忆网络模型在训练集中可有效

捕捉极大值和极小值,较好地反应实测值的极值情况,与循环神经网络、门控循环单元深度学习模型相比,拟合程度高;图 6 给出了训练集确定的双向长短时记忆网络深度学习模型在未知数据测试集上的拟合情况,测试集中双向长短时记忆网络深度学习模型计算值与实测值相比,除个别极值外,其余均与实测值吻合程度高,较好地实现了测试集预测。

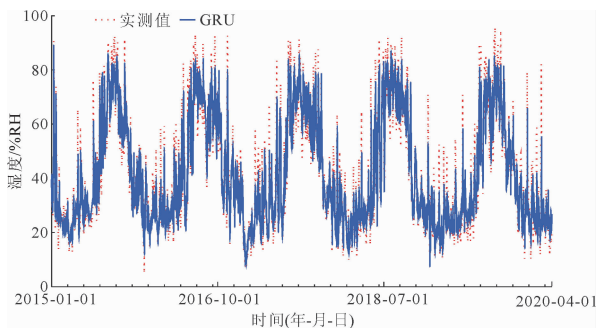


图 3 门控循环单元 (GRU) 深度学习模型训练集计算成果

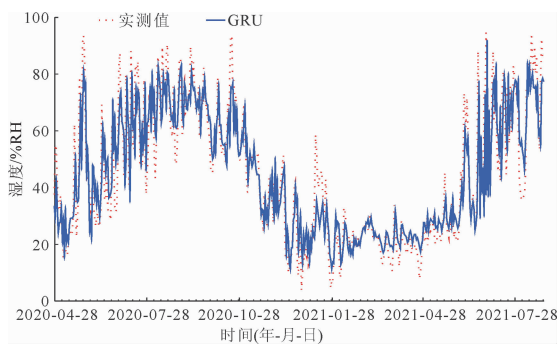


图 4 门控循环单元 (GRU) 深度学习模型测试集计算成果

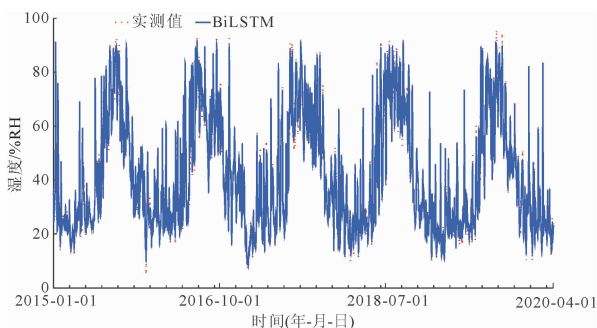


图 5 双向长短时记忆网络 (BiLSTM) 深度学习模型训练集计算成果

图 7—图 8 分别给出了 BiLSTM-Attention 深度学习模型训练集和测试集计算成果,从图 7 中可看出,BiLSTM-Attention 模型在训练集中可有效捕捉极大值和极小值,较好地拟合了实测值的极值分布情况,与循环神经网络、门控循环单元、双向长短时记忆网络深度学习模型相比,拟合程度最高;图 8 给出

了确定的 BiLSTM-Attention 深度学习模型在未知数据测试集中的拟合情况,图 8 表明在测试集中,BiLSTM-Attention 深度学习模型计算值与实测值相比,吻合程度最高,几乎所有极值均得到较好的预测。

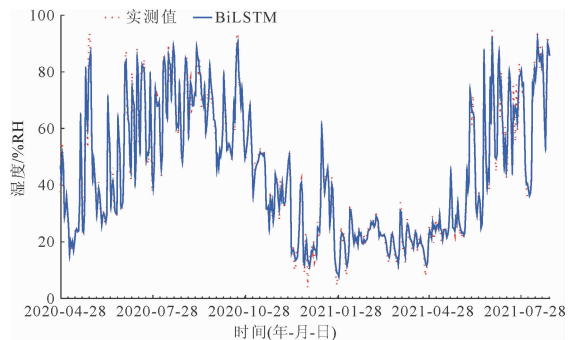


图 6 双向长短时记忆网络 (BiLSTM) 深度学习模型测试集计算成果

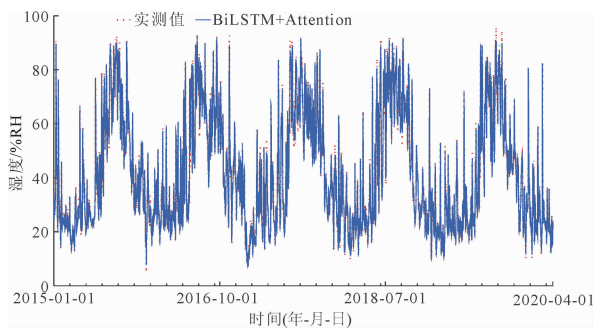


图 7 BiLSTM-Attention 深度学习模型训练集计算成果

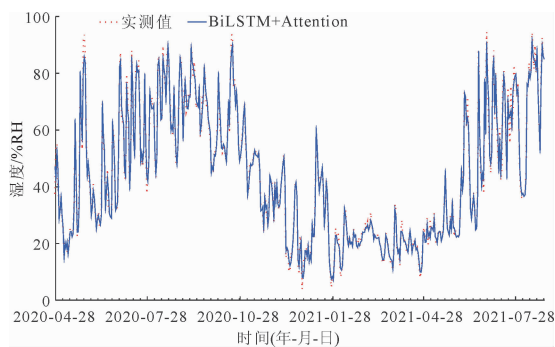


图 8 BiLSTM-Attention 深度学习模型测试集计算成果

对比四种深度学习方法评价指标成果,训练集中  $MAE$ 、 $RMSE$  评价指标门控循环单元模型最大,其次是循环神经网络模型,双向长短时记忆网络模型次之,BiLSTM-Attention 最小,决定系数  $R^2$  规律相反;测试集中  $MAE$ 、 $RMSE$  评价指标循环神经网络模型最大,其次是门控循环单元模型,BiLSTM 模型次之,BiLSTM-Attention 最小,决定系数  $R^2$  规律相反,说明在已知数据进行模型构建时,BiLSTM-Attention、双向长短时记忆网络模型的拟合程度最高,在

对测试集中未知数据进行预测时,模型泛化性也最好,循环神经网络、门控循环单元模型评价指标表明其拟合程度和泛化性较差。决定系数  $R^2$ 、 $MAE$ 、 $RMSE$  评价指标在训练集和测试集中满足一般规律,即测试集评价指标较训练集拟合度略差,在选定的深度学习模型中较好的实现了未知数据的预测。图 1—图 8 绘制的四种不同深度学习模型训练集和测试集的计算成果图中也可看出双向长短时记忆网络、BiLSTM-Attention 两个模型计算结果与实测值吻合良好,BiLSTM-Attention 相对双向长短时记忆网络模型计算结果更贴近湿度实测值。说明采用双向长短时记忆网络、BiLSTM-Attention 深度学习模型对金沙江干热河谷地区进行湿度预测拟合程度高,训练集中预测精度也高,采用双向长短时记忆网络、BiLSTM-Attention 模型进行该地区湿度预测研究是合适的。

## 2.2 模型区间预测结果分析

传统点预测模型仅输出单一预测值,无法量化预测的不确定性。然而,湿度受多种气象因素及局地微气候影响,表现出显著的随机波动特征。尤其在干热河谷地区,由于复杂地形和强烈蒸发作用,湿度变化具有高度非线性与不确定性。因此,仅依赖点预测模型难以满足实际应用需求,亟需发展能够表征预测不确定性的概率预测方法。本研究提出基于 Scott 核密度估计的概率预测框架,通过构建湿度预测值的概率分布,量化未来湿度的可能取值范围及其置信水平。Scott 方法的核心是通过核函数对预测误差分布进行非参数拟合,其概率密度函数可表示为:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (21)$$

式中: $\hat{f}(x)$  为在点  $x$  处的密度估计值; $n$  为样本数量; $x_i$  为第  $i$  个样本数据点; $K$  为核函数(满足非负、对称、积分为 1 的性质); $h$  为带宽。

99%、95% 和 90% 三个置信水平下点预测模型结果较好的 BiLSTM、BiLSTM-Attention 模型的区间预测对比结果:三个置信水平下的预测区间结果呈现出预期趋势,90% 置信水平的区间更窄,而 99% 置信水平的区间更宽。这反映了覆盖概率与区间宽度之间的固有平衡关系——更高的置信水平要求区间必须涵盖更大的不确定性。

为系统评估区间预测性能,本研究采用预测区间覆盖率 (PICP) 和平均区间宽度 (AIW) 作为核心评价指标。其中,预测区间覆盖率表征预测区间的

统计可靠性,为实际观测值落在预测区间内的比例;平均区间宽度测量化预测区间宽度的集中趋势,反映模型对不确定性的表征精度。对计算结果而言,理想的概率预测模型应在保持预测区间覆盖率接近设定置信水平(如 99%、95%、90%)的同时,最小化平均区间宽度值以实现更精确的不确定性量化。

如表 3 所示,通过对比分析两模型的预测区间覆盖率与平均区间宽度指标:当预测区间覆盖率相近时,平均区间宽度较小的模型展现出更优的区间预测性能。以 95% 置信区间结果为例,BiLSTM-Attention 模型在测试集上预测区间覆盖率达到 95.94% (理论值 95%),平均区间宽度为 8.9623% RH,优于 BiLSTM 模型的 PICP = 96.91% (平均区间宽度对应为 17.3764% RH)。这一结果表明,引入注意力机制能有效提升模型对干热河谷湿度随机波动特征的捕捉能力,在保证区间覆盖可靠性的前提下,将预测区间宽度压缩,增强了预测结果的实用价值。整体而言 BiLSTM、BiLSTM-Attention 模型的预测区间覆盖率均接近置信区间水平,两个模型较好地实现了湿度预测,可为金沙江干热河谷地区湿度预测提供决策支持。

表 3 BiLSTM、BiLSTM-Attention 计算模型区间预测成果

计算模型	置信水平 /%	PICP	AIW	计算模型	置信水平 /%	PICP	AIW
BiLSTM	99	0.9938	25.4549	BiLSTM-Attention	99	0.9913	14.3934
BiLSTM	95	0.9691	17.3767	BiLSTM-Attention	95	0.9594	8.9623
BiLSTM	90	0.9527	13.8948	BiLSTM-Attention	90	0.9249	7.1248

## 3 结 论

基于金沙江中游干热河谷地区近 7 年湿度逐日真实值,构建 RNN、GRU、BiLSTM、BiLSTM-Attention 等代理模型进行训练,选取前 80% 数据收作为训练集,将训练后的模型用于剩余 20% 数据进行预测,并通过评价指标对模型的拟合程度,以及对未知数据预测能力进行评判。主要得到以下结论:

(1) 采用 BiLSTM、BiLSTM-Attention 可实现高精度的湿度环境量预测,BiLSTM-Attention 模型计算成果精度最高。

(2) RNN、GRU、BiLSTM、BiLSTM-Attention 四种深度学习模型运行时间分别为:162.16 s、189.82 s、50.00 s、56.20 s;RNN、GRU 两种模型运行时间最长,采用 BiLSTM、BiLSTM-Attention 两种模型时间相差不大,均在 1 min 之内。

(3) BiLSTM、BiLSTM-Attention 两种模型不仅拟合优度高,更展现出优异的泛化能力,其预测精度可满足智慧农业精准灌溉、电力设备防潮预警等实际工程需求。

(4) 基于 Scott 核密度估计的概率预测框架进一步进行了区间预测成果分析;BiLSTM-Attention 模型预测区间覆盖率、平均区间宽度指标均优 BiLSTM 模型,但两模型的预测区间覆盖率均接近置信区间水平,较好地实现了湿度预测,可为金沙江干热河谷地区湿度预测提供决策支持。

#### 参考文献:

- [1] 薛义宽,王禹翔,王 佩,等. 基于 RNN 算法和分子模拟设计新的 STAT3 抑制剂[J]. 中国生物化学与分子生物学报,2024,40(12):1732-1741.
- [2] 吴 昊,李惠仪,庞伟林. 一种基于自适应 RNN 的居民异常用电行为智能检测方法[J]. 人工智能科学与工程,2024(2):69-77.
- [3] 游卉擎,黄鹏程,赵振宇,等. 基于 RNN 的标准单元延时预测方法[J]. 郑州大学学报(理学版),2025,57(3):28-34.
- [4] 王亚茹,杨春旺,屈 卓,等. 双线性特征融合和 GRU 质量聚合的图像质量评价[J/OL]. 智能系统学报,1-12. [2025-08-06]. <https://link.cnki.net/urlid/23.1538.TP.20250806.0912.006>.
- [5] 宁 悦,阿琴花,包雅心,等. 基于门控循环单元和注意力机制的 U 型槽微带贴片天线设计[J]. 激光与光电子学进展,2025,62(15):418-425.
- [6] 柳天虹,齐胜利,裔 扬,等. 基于分位数回归的改进权重 GRU 风电功率区间预测[J]. 太阳能学报,2024,45(12):291-298.
- [7] 刘翌晗,王 艳,马 浩,等. 基于分布式非线性映射和并行输入的 BiLSTM 软测量建模方法[J]. 化工学报,2025,76(7):3373-3387.
- [8] 李鑫帅,贾泽丰,何景源,等. 基于 BiLSTM 的猕猴桃根域土壤水分时序反演方法[J]. 农业工程学报,2025,41(2):112-119.
- [9] 畅 晗,金秀章,赵术善,等. 基于变量筛选和注意力机制的 CNN-BiLSTM 出口 SO<sub>2</sub> 浓度预测模型[J]. 计量学报,2025,46(7):1041-1050.
- [10] 赵法森,刘飞翔,李泽荃,等. 基于 BiLSTM + Attention 模型的煤矿事故隐患自动分类研究[J]. 煤炭科学技术,2022,50(S2):210-217.
- [11] 秦一菲,段珊珊,曹云皓,等. 融合 XGBoost 和 SHAP 的储粮湿度预测及影响因素分析[J/OL]. 中国粮油学报,1-11.
- [12] 罗凤春,付晓雷,蒋晓蕾,等. 基于支持向量机的多层土壤湿度模拟研究[J/OL]. 水文,1-14.
- [13] 张疏桐,张传颂,冯 雄,等. 基于 Kolmogorov-Arnold 网络的日光温室温湿度预测及最小建模数据量[J/OL]. 农业环境科学学报,1-21.
- [14] 毛晓娟,鲍 彤,荀广连,等. 基于 GWO-LSTM 的设施蔬菜温室温度预测[J]. 中国农机化学报,2023,44(1):116-123.
- [15] 智协飞,王 田,季 焱. 基于深度学习的中国地面气温的多模式集成预报研究[J]. 大气科学学报,2020,43(3):435-446.
- [16] 张观山,丁小明,何 芬,等. 基于 LSTM-AT 的温室空气温度预测模型构建[J]. 农业工程学报,2024,40(18):194-201.
- [17] Elman J L. Finding structure in time[J]. Cognitive Science, 1990,14(2):179-211.
- [18] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997,9(8):1735-1780.
- [19] 苏 燕,付家源,林 川,等. 基于时间注意力机制的大坝动态变形预测模型[J]. 水力发电学报,2022,41(7):72-84.
- [20] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate [C]// Proceedings of the 3rd International Conference on Learning Representations (ICLR). 2015.
- [21] 杨国清,刘世林,王德意,等. 基于 Attention-GRU 风速修正和 Stacking 的短期风电功率预测[J]. 太阳能学报,2022,43(12):273-281.